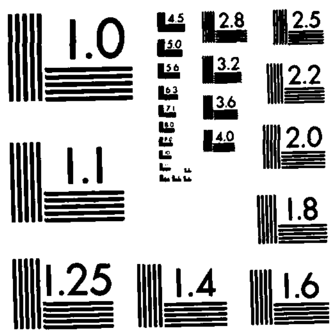
AD-A123 902 SUMS OF FUNCTIONS OF NEAREST NEIGHBOR DISTANCES MOMENT BOUNDS LIMIT THEOR..(U) CALIFORNIA UNIV BERKELEY DEPT OF STATISTICS P J BICKEL ET AL. FEB 82 TR-4 1/1

UNCLASSIFIED N00014-75-C-0444 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SUMS OF FUNCTIONS OF NEAREST NEIGHBOR DISTANCES, MOMENT BOUNDS, LIMIT THEOREMS AND A GOODNESS OF FIT TEST

BY

P. J. BICKEL[1] AND L. BREIMAN[2]

TECHNICAL REPORT NO. 4

FEBRUARY 1982

RESEARCH PARTIALLY SUPPORTED

BY

(1) OFFICE OF NAVAL RESEARCH CONTRACT N00014-75-C-0444
AND N00014-80-C-0163

(2) OFFICE OF NAVAL RESEARCH CONTRACT N00014-82-K-0054

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

SUMS OF FUNCTIONS OF NEAREST NEIGHBOR DISTANCES, MOMENT BOUNDS, LIMIT THEOREMS AND A GOODNESS OF FIT TEST

Peter J. Bickel[1] and Leo Breiman[2]
Department of Statistics
University of California, Berkeley

Summary

We study the limiting behavior of sums of functions of nearest neighbor distances for an m dimensional sample. ~~We establish~~ a central limit theorem and moment bounds for such sums, and an invariance principle for the empirical process of nearest neighbor distances are both established. As a consequence ~~we obtain~~ the asymptotic behavior of a practicable goodness of fit test is obtained based on nearest neighbor distances.

Key Words: Nearest neighbor distances, goodness of fit, asymptotics.

A.M.S. Classification: 60 F 05, 62 G 10.

(1) This research was supported in part by contracts ONR N00014-75-C-0444 and N00014-80-C-0163.

(2) This research was supported in part by ONR Contract N00014-82-K-0054.

1. Introduction and Background

In many areas, there has been a long-standing need for a multidimensional goodness-of-fit test that is general, in the sense that the χ^2 and Kolmogorov-Smirnov test are general in one dimension, and also, is practical in a computational sense. Of course, χ^2 is still available in any number of dimensions, but its usefulness and practicality are virtually nil in high-dimensional spaces.

Take $X_1,\ldots,X_n$ to be n points in m-dimensional Euclidean space selected independently from a distribution with density f(x). Define the nearest neighbor distance R_{jn} from X_j as

$$R_{jn} = \min_{1\leq i\neq j\leq n} \|X_i - X_j\| .$$

In what follows we suppress the dependence of R_{jn} and related quantities on n unless confusion is likely.

The distance d(x,y) between points does not have to be Euclidean. But we assume that it is generated by a norm $\|x\|$, i.e. $d(x,y) = \|x-y\|$.

This paper started with the attempt to derive the limiting distribution of a goodness of fit test for multidimensional densities based on the nearest neighbor distances. We established a form of the invariance principle. Our work had two main byproducts: a central limit theorem for sums of functions of nearest neighbor distances and 4^{th} order moment bounds. These two pieces were then put together to get the invariance result.

The goodness of fit test:

In looking for a practical goodness-of-fit test applicable to densities in an arbitrary number of dimensions, our starting point was the observation, essentially contained in the work by Loftsgaarden and Quesenberry (1965) that the variables

$$U_{jn} = \exp[-n \int_{\|x-X_j\|<R_j} f(x)d\underline{x}] \ , \quad j=1,\ldots,n$$

where $f(x)$ is the underlying density, $X_1,\ldots,X_n$ are n points sampled independently from $f(x)$ and R_j is the distance from X_j to its nearest neighbor, have a univariate distribution that, in any norm $\|\cdot\|$ distance

<u>a</u>; does not depend on $f(x)$

<u>b</u>; is approximately uniform.

The reasoning is simple: let $S(x,r)$ be the sphere with center at x and radius r. For any Borel set A, denote

$$F(A) = \int_A f(y)dy \ .$$

Assume X_1 is the first point selected, then the other n-1. The set $\{R_1 \geq r_1\}$ is equal to the event that none of the $X_2,\ldots,X_n$ fall in the interior of the sphere of radius r_1 about X_1. Hence

$$P(R_1 \geq r_1 | X_1 = x_1) = [1 - F(S(x_1,r_1))]^{n-1} \ .$$

Since for fixed x, $F(S(x,r))$ is monotonically nondecreasing in r, write the above as

$$P[F(S(R_1,x_1)) \geq F(S(r_1,x_1)) | X_1 = x_1] = [1 - F(S(r_1,x_1))]^{n-1} \ .$$

Substituting $z = F(S(x_1,r_1))$ gives

$$P[F(S(x_1,R_1)) \geq z | X_1 = x_1] = (1-z)^{n-1} \qquad (1.1)$$

so that

$$P[F(S(X_1,R_1)) \geq z] = (1-z)^{n-1} \ .$$

Since

$$U_1 = \exp[-nF(S(X_1, R_1))] ,$$

we have that for $\log x > -n$,

$$P(U_1 \le x) = (1 + 1/n \log x)^{n-1} \sim x , \quad \text{for } x \text{ fixed.}$$

The above suggests that a possible approach to a goodness-of-fit test would be to take the density $g(x)$ to be tested, compute the statistics

$$\exp\Big[-n \int_{S(X_j, R_j)} g(x)\, d\underline{x}\Big]$$

and see whether, in some sense, the cumulative distribution function of these n variables is close to the uniform. While this is attractive theoretically, the computations involved in integrating anything but a very simple density over m-dimensional spheres are usually not feasible.

We reasoned that for n large, the nearest neighbor distances were small, on the average, and hence that we could use the approximation

$$\int_{S(X_j, R_j)} g(x)\, dx \sim g(X_j) V(R_j) .$$

where

$$V(r) = K_m r^m$$

is the volume of an m-dimensional sphere of radius r. In this way we were led to testing based on the variables

$$W_j = \exp[-ng(X_j)V(R_j)] , \quad j=1,\ldots,n .$$

An example of a measure of deviation of the W_j variables from the uniform is the statistic

$$S = \sum_1^n (W_{(j)} - j/n)^2$$

where $W_{(j)}$, $j = 1,\ldots,n$, are the ordered W_j variables. Notice that

$$S = n \int_0^1 (\hat{H}(x) - x)^2 d\hat{H}(x)$$

where $\hat{H}(x)$ is the sample d.f. of the W_j.

The invariance principle:

This leads us more generally to studying the stochastic process $\hat{H}(y)$: $0 \leq y \leq 1$, and test statistics based on measures of the deviation of $\hat{H}$ from the uniform or, more appropriately, on the deviations of $\hat{H}$ from its expectation $E\hat{H}$. We had conjectured, based on some simulation studies, that statistics such as S were asymptotically distribution free under the null hypothesis. More generally, we had conjectured that the limiting distribution of $\sqrt{n}(\hat{H}(t) - t)$ was a Gaussian process with zero mean and a covariance not depending on $f(x)$. Our main result, as given in Section 5, is that this is almost true. What holds is that for the sequence of processes

$$Z_n(t) = \sqrt{n}(\hat{H}(t) - E\hat{H}(t))$$
$$Z_n \xrightarrow{w} Z$$

where $Z(t)$, $0 \leq t \leq 1$, is a zero mean Gaussian process whose covariance depends on the hypothesized density g and true density f, and indeed if g = f, then the covariance does not depend on f. The proof of this theorem and other results related to the goodness-of-fit test are given in Section 5.

Defining variables D_{jn} by

$$D_{jn} = n^{1/m} R_{jn} ,$$

then W_{jn} has the form

$$W_j = \phi(X_j, D_{jn})$$

and, denoting the indicator function by $I(\cdot)$,

$$Z_n(t) = \sqrt{n}(\hat{F}(t) - E\hat{F}(t)) = \frac{1}{\sqrt{n}} \sum_1^n [I(W_j \le t) - EI(W_j \le t)]$$

$$= \frac{1}{\sqrt{n}} \sum_1^n [h(X_j, D_j) - Eh(X_j, D_j)]$$

for an appropriate h.

This identification suggests that the appropriate tools for the invariance principle are a central limit theorem and moment bounds and convergence theorems for sums of functions of nearest neighbor distances.

A central limit theorem:

The central limit result established in Sections 3 and 4 is that for a function $h(x,d)$ on $E^{(m)} X [0,\infty) \to E^{(1)}$ such that h is uniformly bounded and almost everywhere continuous with respect to Lebesgue measure,

$$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_1^n h(X_j, D_j)\right) \to \sigma^2 < \infty$$

and

$$\frac{1}{\sqrt{n}} \sum_1^n h^*(X_j, D_j) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$$

where we make the convention here and through the rest of the paper that for any function $h(X_j,D_j)$

$$h^*(X_j,D_j) = h(X_j,D_j) - Eh(X_j,D_j) \quad .$$

This is generalized to a multidimensional central limit theorem, and used to give the result that

$$(Z_n(t_1),\ldots,Z_n(t_k)) \xrightarrow{\mathcal{D}} (Z(t_1),\ldots,Z(t_k)) \quad .$$

Our proof is long. We believe that this is due to the complexity of the problem. Nearest neighbor distances are not independent. But for large sample size the nearest neighbor distance to a point in one region of space is "almost" independent of the nearest neighbor distances in another region of space. The main idea for capitalizing on this large scale independence is to cut the space into a finite number of cells. For any point in a given cell, let its revised nearest neighbor distance be defined using only its neighbors in the <u>same</u> cell. The first step, then, is to show that asymptotically the revised nearest neighbor distances can be substituted for the original nearest neighbor distances. Now, given the number of points in each cell, the set of interpoint distances within the J^{th} cell is independent of those within any other cell. Therefore, given the total cell populations, any sum of functions of the revised nearest neighbor distances is a sum of independent components, with each such component being the sum of the functions of the nearest neighbor distances within a particular cell.

However, the multinomial fluctuation of the cell population is <u>not</u> asymptotically negligible. Thus, the limiting distribution breaks into a sum of two parts, one being the nearly normal sum of the independent cell components given the expected value of the cell populations. The other is an asymptotically normal contribution due to the fluctuations of the cell populations from their expected values. The limiting form of the variance reflects the nature of the problem. It has one term that would be the variance if all nearest neighbor distances were assumed independent. Then there are a number of other, more complex, terms arising from the local dependence.

<u>A moment bound</u>:

Both the central limit theorem and the tightness argument required for the invariance proof rely on moment bounds. Again, there is some difficulty in untangling the dependence between nearest neighbor distances and proving bounds of the type required.

For example, we show in Section 2 that for any measurable function h on $E^{(m)} x[0,\infty) \to E^{(1)}$ with

$$||h|| = \sup|h(x,d)| < \infty$$

there is a constant $M < \infty$ depending only, in a specified and useful way, on h and the dimension m such that

$$E\left(\sum h^{*}(X_j, D_j)\right)^4 \leq Mn^2 \quad .$$

Both the central limit theorem and the moment inequalities (which improve results in Rogers (1977)) should prove generally useful in methods employing nearest neighbor distances.

The plan of the presentation is

Section 2: moment bounds

Section 3: 2^{nd} moment convergence

Section 4: central limit theorem

Section 5: invariance and the goodness-of-fit test

Appendix: technical results on nearest neighbor distances

Section 2 on moment bounds is long and somewhat complex. But the results are needed in the later proofs. The main results of statistical interest are in Sections 4 and 5.

Assumptions on the densities:

Our general assumptions on the density $f(x)$ are that it be uniformly bounded and continuous on its support. These requirements can probably be weakened, but the price may not be worth the extra generality. The following conditions are listed to make the requirements formal.

A: We can choose a version of f such that

(i) $\{f>0\}$ is open

(ii) f is continuous on $\{f>0\}$

(iii) f is uniformly bounded.

Corresponding to A we have:

B: The given function g is nonnegative and

(i) $\{g>0\} \supset \{f>0\}$

(ii) g is continuous on $\{f>0\}$.

Clearly essentially all situations of interest are covered by A and B.

2. Some Useful Moment Inequalities

The central result of this section is the 4^{th} order moment bound (2.2) which is used to prove tightness via Corollary 2.5. We believe it will prove generally useful in the study of procedures based on nearest neighbors. Its formulation and spirit owe much to the excellent thesis of W. R. Rogers (1977). Our method of proof is, however, different from his and suited to the rather delicate estimates we must make.

The proof of the central limit theorem requires only the use of the 2^{nd} order moment bounds given in Lemma 2.11 and its Corollary 2.15. The proofs of 2.11 and 2.15 are given early in this section and the reader interested only in the central limit problem may wish to skip the rest of the section.

The following notation is used:

P is the probability measure making $X_1,\ldots,X_n$ i.i.d. with common density f.

E without subscript is expectation under P.

R_i is the nearest neighbor distance to X_i.

J_i is the index of the nearest neighbor point to X_i.

$D_i = n^{1/m} R_i$

$I(A)$ is the indicator of an event.

$F(A) = \int_A f(y)dy$

$S(x,r) = \{y;\ \|y-x\| \le r\}$

$S_i = S(X_i,R_i)$

For h a measurable function on $E^{(m)} \times [0,\infty) \to E^{(1)}$, denote

$\|h\| = \sup_{x,d} |h(x,d)|$

$h_i = h(X_i,D_i)$

$h_i^* = h_i - Eh_i$

Throughout this section M, with or without a subscript, denotes a finite generic constant depending only on the dimension m.

Theorem 2.1: If $\|h\| < \infty$, then

$$(2.2)\qquad E(\textstyle\sum_i h_i^*)^4 \le Mn^2\|h\|^2[E^2|h_1| + n^4E^2|h_1|F^2(S_1) + n^{-1}\|h\|^2] .$$

Before giving the proof of the theorem we give two corollaries.

Corollary 2.3: Suppose u and w are bounded functions and

$$h(x,d) = u(x)w(x,d) .$$

Then there is a constant $C < \infty$ depending on $\|u\|$, $\|w\|$, m such that

$$(2.4)\qquad E(\textstyle\sum_{i=1}^n h_i^*)^4 \le C(n^2E^2|u(X_1)| + n) .$$

Proof: The corollary follows from

$$E|h_1| \le \|w\|E|u(X_1)|$$

$$E|h_1|F^2(S_1) \le \|w\|E\{E|u(X_1)|E(F^2(S_1))\,|X_1\} = \|w\|E|u(X_1)|\frac{2}{n(n+1)}$$

where the last equality follows from (1.1).

Corollary 2.5: If

$$h(x,d) = I(a \le g(x)d^m \le b)$$

then

$$(2.6)\qquad E(\textstyle\sum_i h_i^*)^4 \le M\{n^2(G_n(b) - G_n(a))^2 + n\}$$

where $G_n(y)$, $y \ge 0$, is the distribution function defined by

$$G_n(y) = (1 - \exp(-\tfrac{n}{2}))^{-1}\int f(x)\left\{1 - \exp[-\tfrac{n}{2}F(S(x,(y/ng(x))^{1/m}))]\right\}dx$$

Proof: Let

$$\alpha(x) = F(S(x,(\frac{a}{ng(x)})^{1/m}))$$

$$\beta(x) = F(S(x,(\frac{b}{ng(x)})^{1/m}))$$

Then, for $j \geq 0$, defining $p_\alpha = F(S(x,\alpha))$, $p_\beta = F(S(x,\beta))$,

$$\begin{aligned} E(|h_1|F^j(S_1)|X_1 = x) &= E[F^j(S(x,R_1)I(p_\alpha \leq F(S(x,R_1)) \leq p_\beta)|X_1 = x] . \\ &= \int_{p_\alpha}^{p_\beta} u^j(n-1)(1-u)^{n-2}du \\ &\leq Mn^{-j} \int_{np_\alpha}^{np_\beta} w^j(1-\frac{w}{n})^{n-2}dw \end{aligned}$$

or

$$\text{(2.7)} \qquad E(|h_1|F^j(S_1)|X_1 = x) \leq M_j n^{-j}(\exp(-\frac{np_\alpha}{2}) - \exp(-\frac{np_\beta}{2})) .$$

If we now apply Theorem 2.1 and use (2.7) for $j = 0,1$ the lemma follows.

The proof of Theorem 2.1 proceeds by a construction similar to one used by Rogers and a series of lemmas.

We assume that we are given a measurable set $S \subset R^m$, $F(S) < 1$, and a set of $r < n$ points, $\underset{\sim}{x} = (x_1,\ldots,x_r)$, where the x_i are fixed points in X. Let $Q_r(\cdot|S,\underset{\sim}{x})$ be the probability measure on $(R^m)^n$ such that $X_1,\ldots,X_{n-r}$ are independent identically distributed with their common distribution being the conditional distribution $F(\cdot|S^c)$ and $X_{n-r+1} = x_i$, $i = 1,\ldots,r$. We write $F(\cdot|S^c)$ as F_S. Its density is, of course,

$$\begin{aligned} f_S(x) &= f(x)/F(S^c) , \quad && x \in S^c \\ &= 0 && \text{otherwise.} \end{aligned}$$

We typically write Q_r for $Q_r(\cdot|S,\underset{\sim}{x})$, and E_{Q_r} to denote the expectation under Q_r.

On a common probability space take $X_1,\dots,X_n$ i.i.d. F and $Y_1,\dots,Y_n$ i.i.d. $F(\cdot|S^c)$ and independent of the X_i and define,

$$\begin{aligned}\tilde{X}_i &= X_i \text{ if } i=1,\dots,n-r \text{ and } X_i \varepsilon S^c\\ &= Y_i \text{ if } i=1,\dots,n-r \text{ and } X_i \varepsilon S\\ &= x_{i-n+r} \text{ if } i=n-r+1,\dots,n\end{aligned}$$

Clearly $\tilde{X}_1,\dots,\tilde{X}_n$ have joint distribution Q_r. Let $\tilde{R}_i$ be the nearest neighbor distance of $\tilde{X}_i$ in the set $\tilde{X}_1,\dots,\tilde{X}_n$ and $\tilde{D}_i$, $\tilde{J}_i$, $\tilde{S}_i$ be defined similarly.

<u>Lemma 2.8</u>: For $n \geq r$, there is a constant M_o such that

$$|E_{Q_r} h(X_1,D_1) - E\, h(X_1,D_1)| \leq ||h||M_o(\frac{r}{n} + F(S))$$

<u>Proof</u>: For $r \geq n/2$, the bound holds trivially. For $n/2 > r$,

(2.9)
$$\begin{aligned}&|E_{Q_r} h(X_1,D_1) - E\, h(X_1,D_1)| =\\ &(n-r)^{-1}\Big|\sum_{i=1}^{n-r} [E\, h(X_i,D_i) - E\, h(\tilde{X}_i,\tilde{D}_i)]\Big|\\ &\leq (n-r)^{-1}\, E \sum_{i=1}^{n-r} |h(X_i,D_i) - h(\tilde{X}_i,\tilde{D}_i)|\\ &\leq (n-r)^{-1}\, ||h||E\sum_{i=1}^{n-r} \{I(X_i \neq \tilde{X}_i) + I(X_i=\tilde{X}_i,\ R_i \neq \tilde{R}_i)\}\end{aligned}$$

Let

$$N = \sum_{i=1}^{n-r} I(X_i \neq \tilde{X}_i)$$

the number of "changed" points among the first $n-r$. Note that $EN = (n-r)F(S)$.
Now

$$I(R_i \neq \tilde{R}_i,\ X_i = \tilde{X}_i) \leq \sum_{j,k} I(J_i = j,\ \tilde{J}_i = k, X_j \neq \tilde{X}_j \text{ or } X_k \neq \tilde{X}_k)$$

and hence

$$\begin{aligned}(2.10)\qquad \sum_i I(R_i \neq \tilde{R}_i,\ X_i = \tilde{X}_i) &\leq \sum_j I(X_j \neq \tilde{X}_j) \sum_i I(J_i = j) \\ &+ \sum_k I(X_k \neq \tilde{X}_k) \sum_k I(J_i = k) \\ &\leq 2\alpha(m)(N+r)\end{aligned}$$

by corollary S1 of the appendix.

From (2.9) - (2.10) and the boundedness of h,

$$|E_{Q_r} h_1 - E\, h_1| \leq ||h|| \left\{(1+2\alpha(m))F(S) + 2\alpha(m)\left(\frac{r}{n-r}\right)\right\}$$

$$\leq ||h|| 2(1+2\alpha(m))(F(S)+\frac{r}{n})$$

and the lemma is proved.

Lemma 2.11: For $\|g\|$, $\|h\| < \infty$, denote $h_1 = h(X_1,D_1)$, $g_2 = g(X_2,D_2)$. Then for $n \geq 4$,

$$|\mathrm{cov}(h_1,g_2)| \leq M_1\|g\|(n^{-1}E|h_1| + E|h_1F(S_1)|) .$$

Proof: Write

$$|\mathrm{cov}(h_1,g_2)| \leq \int_{[J_1=2]} |h_1^* g_2^*| dP + \left|\int_{[J_1\neq 2]} h_1^* g_2^* dP\right| .$$

But

$$\int_{[J_1=2]} |h_1^* g_2^*| dP \leq \frac{2\|g\|}{n-1} \sum_{k=2}^{n} \int_{[J_1=k]} |h_1^*| \leq \frac{4\|g\|}{n-1} E|h_1| . \tag{2.12}$$

Moreover,

$$\int_{[J_1\neq 2]} h_1^* g_2^* dP = \int_{[J_1\neq 2]} h_1^* \{E(g_2|X_1,X_{J_1},J_1) - Eg_2\} dP . \tag{2.13}$$

On the set $J_1 \neq 2$, given $X_1 = x_1$, $X_{J_1} = x_2$, the $\{X_j,\ 2 \leq j \leq n,\ j \neq J_1;\ X_1, X_{J_1}\}$ are distributed according to $Q_2(\cdot|S(x_1,|x_2-x_1|),(x_1,x_2))$. By Lemma 2.8

$$\begin{aligned} \left|\int_{[J_1\neq 2]} h_1^* g_2^* dP\right| &\leq \int_{[J_1\neq 2]} |h_1^*| M_0\|g\|(2n^{-1} + F(S_1)) dP \\ &\leq 4M_0\|g\|[n^{-1}E|h_1| + E|h_1F(S_1)|] \end{aligned} \tag{2.14}$$

and the lemma follows from (2.12)-(2.14).

Corollary 2.15: For $\|h\|$, $\|g\| < \infty$, and for $n \geq 4$,

$$|\mathrm{cov}(h_1,g_2)| \leq M_2\|g\|(Eh_1^2)^{1/2}/n .$$

Proof: From (1.1) it follows that $EF^2(S_1) = 2/n(n+1)$. Now apply the Schwartz inequality.

The bounds in Lemma 2.11 and Corollary 2.15 can clearly be made symmetric in h_1 and g_2. We use them primarily for

<u>Lemma 2.16</u>: $$|\mathrm{cov}_{Q_r}(h_1,h_2) - \mathrm{cov}(h_1,h_2)| \le ||h||^2 M_3\left(\frac{r^2}{n^2} + F^2(S)\right)$$

<u>Proof</u>: Let $(X_1',\tilde{X}_1'),\ldots,(X_n',\tilde{X}_n')$ have the same joint distribution as the vector $\{(X_1,\tilde{X}_1),\ldots,(X_n,\tilde{X}_n)\}$ and be independent of that vector. Let primes on $D_i,\tilde{D}_i,J_i$, etc. as usual denote calculations based on the appropriate sample. Then

$$(2.17)\qquad \mathrm{cov}(h_1,h_2) - \mathrm{cov}_{Q_r}(h_1,h_2) = \frac{1}{2} E\,\Delta$$

$$\Delta = (h_1-h_1')(h_2-h_2') - (\tilde{h}_1-\tilde{h}_1')(\tilde{h}_2-\tilde{h}_2')$$

where

$$h_i' = h(X_i',D_i'),\ \tilde{h}_i = h(\tilde{X}_i,\tilde{D}_i),\ \tilde{h}_i' = h(\tilde{X}_i',\tilde{D}_i')$$

The proof proceeds by a series of steps.

Let

$$E_i = \{h_i \ne \tilde{h}_i\}$$

$$E_i' = \{h_i' \ne \tilde{h}_i'\}$$

Since

$$I(E_i) \le I(X_i \ne \tilde{X}_i) + I(X_i = \tilde{X}_i,\ R_i \ne \tilde{R}_i)$$

Lemma A.1 and elementary arguments yield that

(2.18) $\max\{P(E_i \cap E_j),\ P(E_i \cap E_k') : \text{all } i,j,k, i \neq j\} \leq M\left(\frac{r^2}{n^2} + F^2(S)\right)$

Since $\Delta = 0$ on $[\cup_{i=1}^{2}\{E_i \cup E_i'\}]^c$, (2.18) and symmetry arguments imply that

(2.19)
$$|E\Delta| \leq 4|E(h_1-\tilde{h}_1)(h_2-h_2)I(E_1E_2^c[E_1']^c[E_2']^c)|$$
$$+ M\|h\|^2\left(\frac{r^2}{n^2} + F^2(S)\right)$$

Using lemma A.1 again we bound the first term on the right hand side of (2.19) by,

(2.20)
$$4|E\{(h_1-\tilde{h}_1)(h_2-h_2')(I(J_1\neq 2,\tilde{J}_1\neq 2,X_2=\tilde{X}_2)[I(X_1\neq\tilde{X}_1)+I(X_1=\tilde{X}_1,\ R_1\neq\tilde{R}_1)]\}|$$
$$+ M\|h\|^2\left(\frac{r^2}{n^2} + F^2(S)\right)$$

Let $\Xi = \{i : X_i \neq \tilde{X}_i\}$. Given Ξ, $X_i, i\in\Xi, X_1, X_{J_1}, \tilde{X}_1, X_{\tilde{J}_1}, \tilde{X}_{\tilde{J}_1}$ and $X_2=\tilde{X}_2$ the variables $X_1,\ldots,X_n$ can be permuted to have a

$$Q_r(\cdot|S(X_1,R_1)\cup S(\tilde{X}_1,\tilde{R}_1),\{X_i,\ i\in\Xi,\ X_1,X_{J_1},X_{\tilde{J}_1}\})$$

distribution with X_2 in the lead and $r = N+I(X_1=\tilde{X}_1) + I(X_{J_1}=\tilde{X}_{J_1}) + I(X_{\tilde{J}_1}=\tilde{X}_{\tilde{J}_1})$.

Conditioning on this information within the expectation in (2.20) and using the independence of h_2' we can apply lemma 2.8 to the difference between the conditional expectation of h_2 and Eh_2' and bound the first term in (2.20) by

$$(2.21)\qquad 4||h||^2 M_0(m)\, E\left\{\left(I(X_1\neq\tilde{X}_1) + I(X_1=\tilde{X}_1, R_1\neq\tilde{R}_1)\right)\left(\frac{N+3}{n} + F(S_1) + F(\tilde{S}_1)\right)\right\}$$

Estimates of the order $\frac{r^2}{n^2} + F^2(S)$ for all the terms in (2.21) are given in lemma A.2. Combining (2.19) - (2.21) the lemma follows.

<u>Lemma 2.22</u>:

$$(2.23)\qquad \left|Eh_1^*h_2^*h_3^*h_4^*\right| \leq M_4\, ||h||^2\left(\frac{E^2|h_1|}{n^2} + n^2E^2|h_1|F^2(S_1) + ||h||^2\, n^{-3}\right)$$

<u>Proof</u>: Let $E_{12} = [J_1, J_2 \notin \{3,4\}]$, $\pi = h_1^*h_2^*h_3^*h_4^*$

Then,

$$(2.24)\qquad \int_{E_{12}} \pi\, dP = \int_{E_{12}} h_1^*h_2^*\, \{\mathrm{cov}_{Q_r}(h_1,h_2) + (E_{Q_r}h_1 - Eh_1)^2\}dP$$

where

$$Q_r = Q_r(\cdot|S(X_1,R_1)US(X_2,R_2), \{X_1,X_2,X_{J_1},X_{J_2}\})\text{ and } r \leq 4.$$

Apply lemmas 2.8, 2.11 and 2.16 to get,

$$(2.25)\qquad \left|\int_{E_{12}} \pi\, dP\right| \le \left(M_1 \|h\| (n^{-1}E|h_1| + E|h_1F(S_1)|)\right) \times \left|\int_{E_{12}} h_1^* h_2^* dP\right|$$

$$+ M_2 \|h\|^2 \int_{E_{12}} |h_1^* h_2^*| (n^{-2} + F^2(S_1) + F^2(S_2))dP$$

Next

$$(2.26)\qquad \int_{E_{12}^c} h_1^* h_2^* = 2\int_{[J_1=3]} h_1^* h_2^*$$

$$+ 2\int_{[J_2=3, J_1 \notin \{3,4\}]} h_1^* h_2^*$$

Condition in the first integral on the right in (2.26) by X_1, X_{J_1}, J_1 and apply lemma 2.8 to get the bound

$$(2.27)\qquad 2M_0 \|h\| \int_{[J_1=3]} |h_1^*| (n^{-1} + F(S_1))dP$$

$$\le 4M_0 \frac{\|h\|}{n-1} (n^{-1}E|h_1| + E|h_1F(S_1)|)$$

by the usual symmetry argument. Condition in the second integral by X_2, X_{J_2}, J_2 and obtain a bound as in (2.27). Conclude that

$$\left|\int_{E_{12}} h_1^* h_2^*\right| \le |\mathrm{cov}(h_1, h_2)| + M \frac{\|h\|}{n} (E|h_1| n^{-1} + E|h_1F(S_1)|)$$

and hence that the first term in (2.25) is bounded by

$$(2.28) \qquad M\,\|h\|^2\left(\frac{E^2|h_1|}{n^2} + E^2|h_1 F(S_1)|\right) \quad .$$

On the other hand, applying lemma 2.8 again

$$(2.29) \qquad \int |h_1^* h_2^*|(n^{-2} + F^2(S_1)) \le \|h\| \int_{[J_1=2]} |h_1^*|(n^{-2} + F^2(S_1))$$

$$+ \int_{[J_1 \neq 2]} |h_1^*|(n^{-2} + F^2(S_1))\{E|h_2^*| + M_0\,\|h\|\,(n^{-1} + F(S_1))\} \quad .$$

The first term in (2.29) is $\le M\,\|h\|^2 n^{-3}$ by the usual symmetry argument. The second is

$$\le M(E^2|h_1|n^{-2} + E|h_1|\,E|h_1|F^2(S_1)) + \|h\|^2 n^{-3})$$

$$(2.30) \qquad \le M(2(E^2|h_1|n^{-2} + n^2\,E^2|h_1|F^2(S_1)) + \|h\|^2\,n^{-3})$$

and hence combining (2.28) and (2.30) we get

$$(2.31) \qquad |\int_{E_{12}} \pi\,dP| \le M\,\|h\|^2\left(\frac{E^2|h_1|}{n^2} + E^2|h_1|F(S_1)\right.$$

$$\left. + n^2\,E^2|h_1|F^2(S_1) + \|h\|^2\,n^{-3}\right) \quad .$$

Now consider

$$(2.32) \qquad \int_{[J_1=3]} \pi\,dP = \int_{[J_1=3,J_3\notin\{2,4\}]} \pi\,dP + 2\int_{[J_1=3,J_3=2]} \pi\,dP$$

By conditioning on $X_1, X_3, J_1, J_3, X_{J_1}, X_{J_3}$ we can bound the first integral on the right in (2.32) in exactly the same way as $\int_{E_{12}} \pi \, dP$ by,

$$(2.33) \qquad M\left\{\left(||h||\left(\frac{E|h_1|}{n} + E|h_1 F(S_1)|\right)\right| \int_{[J_1=3, J_3\notin\{2,4\}]} h_1^* h_3^* dP| \right.$$

$$\left. + ||h||^2 \int_{[J_1=3]} |h_1^* h_3^*| (n^{-2} + F^2(S_1) + F^2(S_3)) dP \right\}$$

Now use symmetry to bound

$$\left| \int_{[J_1=3, J_3\notin\{2,4\}]} h_1^* h_3^* \right|$$

by

$$\frac{2||h||}{n-1} E|h_1|$$

and the second term in (2.33) by,

$$\frac{M||h||^4}{n^3} .$$

Hence,

$$(2.34) \qquad \left| \int_{[J_1=3, J_3\notin\{2,4\}]} \pi \, dP \right| \le M||h||^2 \left(\frac{E^2|h_1|}{n^2} + E^2|h_1| F(S_1) \right.$$

$$\left. + ||h||^2 n^{-3} \right)$$

Next write,

$$(2.35)\qquad \int_{[J_1=3,J_3=2]} \pi\, dP = \int_{[J_1=3,J_3=2,J_2\neq 4]} \pi\, dP + \int_{[J_1=3,J_3=2,J_2=4]} \pi\, dP$$

Now

$$(2.36)\qquad P[J_1=3,J_3=2,J_2=4] = \frac{1}{n-3} \sum_{i=4}^{n} P[J_1=3,J_3=2,J_2=i]$$

$$\leq (n-3)^{-1}\, P[J_1=3,J_3=2] \leq (n-3)^{-1}(n-2)^{-1} P[J_1=3]$$

$$\leq M\, n^{-3}$$

Hence,

$$(2.37)\qquad \left|\int_{[J_1=3,J_3=2,J_2=4]} \pi\, dP\right| \leq M||h||^4\, n^{-3}$$

Next condition on $X_1, X_2, X_3, J_1, J_2, J_3, R_1, R_2, R_3$ in the first term of (2.35) and apply lemma 2.8 to get

$$(2.38)\qquad \left|\int_{[J_1=3,J_3=2,J_2\neq 4]} \pi\, dP\right| \leq M_0||h||^4 \int_{[J_1=3,J_3=2]} (n^{-1} + \sum_{i=1}^{3} F(S_1))\, dP$$

Now,

$$P[J_1=3,\ J_3=2] \le Mn^{-2}$$

as in (2.36) and similarly,

(2.39) $$\int_{[J_1=3,J_3=2]} F(S_1)dP \le (n-2)^{-1} \int_{[J_1=3]} F(S_1)dP$$

$$= [(n-2)(n-1)]^{-1}\ EF(S_1) \le Mn^{-3}$$

(2.40) $$\int_{[J_1=3,J_3=2]} F(S_2)dP = (n-2)^{-1} \int_{[J_3=2]} F(S_2) \sum_{i\neq 2,3} I(J_i=3)dP$$

$$\le (n-2)^{-1}\alpha(m) \int_{[J_3=2]} F(S_2)dP$$

by corollary S1,

$$\le [(n-2)(n-1)]^{-1}\alpha^2(m) \int F(S_2)dP \le Mn^{-3}$$

(2.41) $$\int_{[J_1=3,J_3=2]} F(S_3)dP \le [(n-2)(n-1)]^{-1}\alpha(m)EF(S_3) \le Mn^{-3}$$

Combining these estimates with (2.38), (2.37) and (2.35) we get,

(2.42) $$\left|\int_{[J_1=3,J_3=2]} \pi\, dP\right| \le M||h||^4 n^{-3}$$

and hence from (2.32), (2.34) and (2.42),

$$(2.43)\qquad \left|\int_{[J_1=3]} \pi\, dP\right| \le M\|h\|^2 \left(\frac{E^2|h_1|}{n^2} + E^2|h_1|F(S_1) + \|h\|^2 n^{-3} \right)$$

Next consider,

$$(2.44)\qquad \int_{[J_2=3,J_1\notin\{3,4\}]} \pi\, dP = \int_{[J_2=3]} \pi\, dP - \int_{[J_1=J_2=3]} \pi\, dP - \int_{[J_2=3,J_1=4]} \pi\, dP$$

Of these terms the first is bounded in (2.43). The next is written,

$$(2.45)\qquad \int_{[J_1=J_2=3,\ J_3\neq 4]} \pi\, dP + \int_{[J_1=J_2=3,\ J_3=4]} \pi\, dP\ .$$

The second term in (2.45) is bounded by $M\|h\|^4 n^{-3}$ as in (2.40). The first (conditioning on X_1, X_2, X_3, etc.) is bounded by

$$M\|h\|^4 \int_{[J_1=J_2=3]} (n^{-1} + \sum_{i=1}^{3} F(S_i))dP$$

and again by $M\|h\|^4 n^{-3}$ by arguing as in (2.39) - (2.41). For example,

$$\int_{[J_1=J_2=3]} F(S_1)dP \le \frac{\alpha(m)}{n-2} \int_{[J_1=3]} F(S_1)dP = \alpha(m)[n(n-1)(n-2)]^{-1}$$

Finally,

$$(2.46)\qquad \left|\int_{[J_2=3,J_1=4]} \pi \, dP\right| \le \|h\| \int_{[J_2=3,J_1=4]} |h_1^* h_2^* h_3^*|$$

$$\le (n-3)^{-1} \|h\| \int_{[J_2=3]} |h_1^* h_2^* h_3^*|$$

$$\le [(n-3)(n-2)]^{-1} \|h\|^2 \, E|h_1^* h_2^*|$$

$$\le Mn^{-2} \|h\|^2 \, (E^2|h_1| + \operatorname{cov}(|h_1^*|, |h_2^*|))$$

$$\le Mn^{-2} \|h\|^2 \, (E^2|h_1| + \|h\|^2 \, n^{-1})$$

by lemma 2.11. By our discussion and (2.43) - (2.46),

$$(2.47)\qquad \left|\int_{E_{12}^c} \pi \, dP\right| \le M\|h\|^2 \left(\frac{E^2|h_1|}{n^2} + E^2|h_1|F(S_1) + \|h\|^2 \, n^{-3}\right)$$

Now by the Schwartz inequality,

$$E^2|h_1|F(S_1) \le E|h_1| \; E|h_1|F^2(S_1)$$

$$\le \frac{E^2|h_1|}{n^2} + n^2 E^2|h_1|F^2(S_1)$$

The lemma, therefore, follows from (2.31) and (2.47).

<u>Lemma 2.48</u>: For $M_5 < \infty$

$$|E[h_1^*]^2 h_2^* h_3^*| \leq M_5 ||h||^2 \left(\frac{E^2|h_1|}{n} + nE^2|h_1|F(S_1) + \frac{||h||^2}{n^2} \right) \tag{2.49}$$

<u>Proof</u>: The argument goes much as for lemma 2.22 and is sketched. If we denote the integrand by π^*

$$|\int_{[J_1 \neq 2,3]} \pi^* dP| \leq M||h|| \Big\{ (n^{-1}E|h_1| + E|h_1|F(S_1))$$

$$\times \int_{[J_1 \neq 2,3]} [h_1^*]^2 + ||h||^3 n^{-2} \Big\}$$

$$\leq M||h||^2 \{n^{-1}E^2|h_1| + nE^2|h_1|F(S_1) + ||h||^2 n^{-2}\},$$

while

$$|\int_{[J_1=2]} [h_1^*]^2 h_2^* h_3^* dP| \leq ||h||^2 \int_{[J_1=2]} |h_1^* h_3^*| dP \leq Mn^{-1} ||h||^2 \int |h_1^* h_2^*| dP$$

$$\leq M||h||^2 n^{-1}(E^2|h_1| + n^{-1}||h||^2) \text{ arguing as in (2.46).}$$

The lemma follows.

<u>Proof of Theorem</u>: Write

$$(2.50) \qquad E\left(\sum_i h_i^*\right)^4 \le n\, E[h_1^*]^4 + 6n(n-1)\, E[h_1^*]^2[h_2^*]^2$$

$$+ 6n(n-1)(n-2)\, |E[h_1^*]^2 h_2^* h_3^*| + n(n-1)(n-2)(n-3)\,|E h_1^* h_2^* h_3^* h_4^*|$$

We apply lemmas 2.22 and 2.48 to the last two terms of (2.50); note that the second term is

$$\le 6n^2 ||h||^2 \, (E^2|h_1^*| + |\mathrm{cov}(|h_1^*|,|h_2^*|)|)$$

and apply lemma 2.11, and bound $E[h_1^*]^4$ by $16||h||^4$.

The theorem follows.

3. Second Moment Convergence

The central result of this section is the evaluation of the limit of $\mathrm{Var}(\frac{1}{\sqrt{n}} \sum_1^n h(X_j, D_j))^2$ for a certain class of functions h. Starting with the density f(x), define

$$\gamma(x) = f(x)^{-1/m} ,$$

and for any measurable function h on $E^{(m)} \times [0,\infty) \to E_1$, let

$$\tilde{h}(x,r) = h(x,\gamma(x)r) .$$

Define L_0, L_1, L_2 as functions of bounded variation given by

$$(3.1) \qquad L_0(r) = e^{-V(r)}$$

$$(3.2) \quad L_1(r_1,r_2) = e^{-V(r_1)-V(r_2)}[V(r_1) + V(r_2) - V(r_1)V(r_2)]$$

$$(3.3) \quad L_2(r_1,r_2) = e^{-V(r_1)-V(r_2)}[\int_{B(r_1,r_2)}(e^{V(r_1,r_2,z)} - 1)dz - V(\max(r_1,r_2))]$$

where

$$B(r_1,r_2) = \{z;\ \max(r_1,r_2) \le \|z\| \le r_1+r_2\}$$

$$V(r_1,r_2,z) = \int_{S(0,r_1)\cap S(z,r_2)} dy$$

For any two functions h, h' define the functional L(h,h') by

$$(3.4) \quad L(h,h') = \int \tilde{h}(x_1,r_1)\tilde{h}'(x_2,r_2)f(x_1)f(x_2)L_1(dr_1,dr_2)dx_1dx_2 + \int \tilde{h}(x,r_1)\tilde{h}'(x,r_2)f(x)L_2(dr_1,dr_2)dx$$

The moment convergence result is

<u>Theorem 3.5</u>: If h is measurable on $E^{(m)} \times [0,\infty) \rightarrow E^{(1)}$ and satisfies

(i) $\|h\| < \infty$

(ii) the set of discontinuities of h has Lebesgue measure 0,

then

$$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_1^n h(X_i, D_i)\right) \rightarrow \sigma^2(h)$$

where

$$(3.6) \quad \sigma^2(h) = \int \tilde{h}^2(x,r) f(x) L_0(dr)dx - \left[\int \tilde{h}(x,r) f(x) L_0(dr)dx\right]^2 + L(h,h) \ .$$

As the proof will reveal, the first two terms of (3.6) would be the limit if the R_j were independent. The L(h,h) term is contributed by the local dependence of the nearest neighbor distances.

The proof of the theorem is split into two pieces. Proposition 3.7 below shows that the diagonal terms in

$$\frac{1}{n}\left(\sum_1^n h^*(X_i, D_i)\right)^2$$

converge to the first two terms of (3.6). Then proposition 3.20 gives convergence of the off-diagonal terms to L(h,h). We assume throughout that the conditions of the theorem hold.

Let X, D be a random m vector and nonnegative random variable respectively such that X has density f and

$$P[D > r | X] = \exp\{-f(X)V(r)\} \ .$$

Equivalently, $D/\gamma(X)$ is independent of X and

$$P[D/\gamma(X) > r] = L_0(r) \ .$$

Proposition 3.7: Let f satisfy A(i)-(iii). Then, as $n \to \infty$,

$$(X_1, D_{1n}) \xrightarrow{\mathcal{L}} (X, D)$$

where (X_1, D_{1n}) is used to stand generically for the common law of any of the pairs (X_i, D_i) and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution. Therefore

$$(3.8) \qquad Eh(X_1, D_{1n}) \to \int \tilde{h}(x,r) f(x) L_0(dr)dx$$

$$(3.9) \quad \text{Var } h(X_1, D_{1n}) \to \int \tilde{h}^2(x,r) f(x) L_0(dr)dx - \left(\int \tilde{h}(x,r) f(x) L_0(dr)dx\right)^2 .$$

Proof: Almost immediate, since

$$P(D_{1n} > r \mid X_1 = x) \to e^{-f(x)V(r)} = P(D > r \mid X = x)$$

and the set of discontinuities of h has probability zero with respect to the (X,D) distribution.

Proposition 3.10: For h(x,r) any function satisfying the hypothesis of theorem 3.5

$$n \text{ Cov}(h(X_1, D_1), h(X_2, D_2)) \to L(h,h) .$$

Proof: It is, we assert, sufficient to show for any two functions ϕ_1, ϕ_2 of the form

$$(3.11) \qquad \phi_i(x,r) = g_i(x) I(r \geq r_i) , \quad i = 1, 2$$

with $g_i(x)$ *uniformly continuous and bounded, that*

$$(3.12) \qquad n \text{ Cov}(\phi_1(X_1, D_1), \phi_2(X_2, D_2)) \to L(\phi_1, \phi_2) .$$

To see this note that if $\mathcal{J}$ is the set of all finite linear combinations of functions of the form (3.11) then we can get a sequence $h_k \in \mathcal{J}$ such that

$$\|h_k\| \leq 2\|h\|$$

and with respect to L-measure on $E^{(m)} \times [0,\infty)$, $h_k \to h$ a.e. (since h is a.e. continuous). Now

$$\text{(3.13)} \quad \begin{aligned} &\mathrm{Cov}(h(X_1,D_1),h(X_2,D_2)) - \mathrm{Cov}(h_k(X_1,D_1),h_k(X_2,D_2)) \\ &= \mathrm{Cov}(h(X_1,D_1) - h_k(X_1,D_1), h(X_2,D_2) + h_k(X_2,D_2)) . \end{aligned}$$

Using corollary 2.15 on (3.13) gives the bound

$$\overline{\lim_n}\left|\mathrm{Cov}(h(X_1,D_1),h(X_2,D_2)) - \mathrm{Cov}(h_k(X_1,D_1),h_k(X_2,D_2))\right| \leq c\|h\|(E|h-h_k|^2)^{1/2} .$$

Now the bounded convergence theorem gives $E(h-h_k)^2 \to 0$, and (3.12) implies that

$$\mathrm{Cov}(h_k(X_1,D_1),h_k(X_2,D_2)) \to L(h_k,h_k) .$$

Since $L(h_k,h_k) \to L(h,h)$, the assertion follows.

<u>Proof of (3.12)</u>: For $i = 1,2$, let

$$S_i = S(x_i,n^{-1/m}r_i) , \quad F_i = F(S_i) , \quad F_{12} = F(S_1 \cap S_2)$$

and let

$$A = \{(x_1,x_2);\ \|x_1-x_2\| \geq n^{-1/m}(r_1+r_2)\}$$

$$B = \{(x_1,x_2);\ n^{-1/m}\max(r_1,r_2) \leq \|x_1-x_2\| \leq n^{-1/m}(r_1+r_2)\}$$

$$C = \{(x_1,x_2),\ \|x_1-x_2\| \leq n^{-1/m}\max(r_1,r_2)\} .$$

Then

$$P(R_1 \geq n^{-1/m}r_1,\ R_2 \geq n^{-1/m}r_2 | X_1=x_1,X_2=x_2) = \begin{cases} (1-F_1-F_2)^{n-2} & ,\ (x_1,x_2) \in A \\ (1-F_1-F_2+F_{12})^{n-2}, & (x_1,x_2) \in B \\ 0 & ,\ (x_1,x_2) \in C \end{cases}$$

and

$$P(R_i \geq n^{-\frac{1}{m}} r_i | X_i = x_i) = (1-F_i)^{n-1} \quad .$$

Then, denoting

$$L(x_1,x_2,r_1,r_2) = P(R_1 \geq n^{-\frac{1}{m}} r_1, R_2 \geq n^{-\frac{1}{m}} r_2 | X_1 = x_1, X_2 = x_2) - [(1-F_1)(1-F_2)]^{n-1}$$

and $g_1(x_1)$ by g_i, $f(x_a)$ by f_i,

$$\begin{aligned}
Cov(\phi_1,\phi_2) &= \int g_1(x_1) g_2(x_2) L(x_1,x_2,r_1,r_2) f(x_1) f(x_2) dx_1 dx_2 \\
&= \int g_1 g_2 [(1-F_1-F_2)^{n-2} - (1-F_1)^{n-1}(1-F_2)^{n-1}] f_1 f_2 \\
&+ \int_B g_1 g_2 [(1-F_1-F_2+F_{12})^{n-2} - (1-F_1-F_2)^{n-2}] f_1 f_2 \\
&- \int_C g_1 g_2 [(1-F_1-F_2)^{n-2}] f_1 f_2 \\
&= I_1 + I_2 - I_3 \quad .
\end{aligned}$$

Because $nF_i \leq \bar{f} V(r_i)$, where $\bar{f}$ is the supremum of f, and $nF_i \to f(x_i)V(r_i)$, for fixed x_1, x_2

$$n[(1-F_1-F_2)^{n-2} - (1-F_1)^{n-1}(1-F_2)^{n-1}] =$$

$$n(1-F_1)^{n-2}(1-F_2)^{n-2}\left[\left\{1 - \frac{F_1 F_2}{(1-F_1)(1-F_2)}\right\}^{n-2} - (1-F_1)(1-F_2)\right]$$

$$e^{-f(x_1)V(r_1)-f(x_2)V(r_2)}[f(x_1)V(r_1)+f(x_2)V(r_2)-f(x_1)f(x_2)V(r_1)V(r_2)] \quad .$$

Furthermore, the convergence is bounded. Therefore

$$n\, I_1 \to \int \tilde{\phi}(x_1,r_1)\tilde{\phi}(x_2,r_2)L_1(dr_1,dr_2)f(x_1)f(x_2)dx_1dx_2$$

as can be seen by making the transformations $V(r_i') = f(x_i)V(r_i)$.

In I_2, I_3 make the transformation

$$x_2 = x_1 + n^{-\frac{1}{m}}z \quad ,$$

leading to

$$B = \{(x_1,z);\ \max(r_1,r_2) \leq \|z\| \leq r_1 + r_2\}$$

$$C = \{(x_1,z);\ \|z\| \leq \max(r_1,r_2)\} \quad .$$

On BUC, for x_1 fixed

$$f(x_2)g_2(x_2) \to f(x_1)g_2(x_1)$$

uniformly, and

$$n\, F_i \to f(x_1)V(r_i)$$

$$n\, F_{12} \to f(x_1)V(r_1,r_2,z)$$

where

$$V(r_1,r_2,z) = \int_{\|y\| \leq r_1,\ \|y-z\| \leq r_2} dy$$

Therefore

$$nI_2 \longrightarrow \int [\int_B (e^{f(x)V(r_1,r_2,z)} - 1)dz]e^{-f(x)[V(r_1+V(r_2)]} g_1(x)g_2(x)f^2(x)dx \ .$$

A simpler argument gives

$$nI_3 \longrightarrow \int V(\max(r_1,r_2))e^{-f(x)[V(r_1)+V(r_2)]} g_1(x)g_2(x)f^2(x)dx \ .$$

In both integrals, make the substitution $V(r_i') = f(x)V(r_i)$ and add the limits together to get the proposition.

4. A Central Limit Theorem

The main result of this section is

Theorem 4.1: Suppose the set of discontinuities of h has Lebesgue measure 0 in $E^{(m)} \times [0,\infty)$ and

$$\sup_{x,d} |h| = \|h\| < \infty .$$

Then if the density of the distribution satisfies A(i)-(iii),

$$\frac{1}{\sqrt{n}} \sum_1^n h^*(X_j, D_j) \xrightarrow{\mathcal{L}} N(0, \sigma^2(h)) \tag{4.2}$$

where $\sigma^2(h)$ is given in Theorem 3.5.

The proof proceeds in a series of propositions.

Notational convention:

Lower case c denotes a constant depending only on m and $\|h\|$. The dependence of other constants on various auxiliary parameters introduced below will be noted as needed.

Proposition 4.3: There exists a sequence of bounded sets $C_N \subset E^{(m)}$ with $C_N \subset C_{N+1}$ such that

1) $\text{diameter}(C_N) \le N$

2) $\inf_{x \in C_N} f(x) = \delta_N > 0$

3) $P(X \in C_N^c) \to 0$

Proof: There exist compact sets $A_N \subset A_{N+1}$ such that $\int_{A_N} f\,dx \to 1$. Choose $\delta_N > 0$ such that $\delta_N \int_{A_N} dx \to 0$. Let

$$F_N = \{x;\ f(x) \ge \delta_N\}$$

and take $C_N = A_N \cap F_N$. Then

$$\int_{A_N} f - \int_{C_N} f \le \int_{A_N \cap F_N^c} f \le \delta_N \int_{A_N} dx$$

so $\int_{C_N} f \to 1$.

In preparation for the next step, let D_N be a cube of side N such that $C_N \subset D_N$. Divide D_N into $L = (k)^m$ congruent subcubes $D_{N,\ell}$, $\ell = 1,\ldots,L$, and let

$$B_\ell = \bar{D}_{N,\ell} \cap C_N \ , \quad \ell = 1,\ldots,L$$

$$\tilde{B} = \bigcup_\ell \partial(B_\ell)$$

where ∂ denotes boundary. The B_ℓ, $\ell = 1,\ldots,L$ provide the basic cells such that nearest neighbor links between different cells will be cut. _From now on until the end of the string of propositions_ N _and the_ B_ℓ, $\ell = 1,\ldots,L$ _will be fixed_.

Select $d_N > 0$ and let

$$E_N = \{x;\ x \in C_N,\ d(x,\tilde{B}) \ge d_N\}$$

where $d(x,\tilde{B})$ is the distance from x to the set $\tilde{B}$. Write (X,D) for (X_1, D_{1n}). Note that by using $f(x) \le \sup_x f(x) = \bar{f}$, we get

$$P(X \in C_N,\ d(X,\tilde{B}) < d_N) \le 2md_N L^{1/m} N^{m-1} \bar{f} \ .$$

Now let

$$\underline{h}(x,d) = I(x \in E_N) h(x,d) \ .$$

We suppress dependence on N, L here and in the sequel except where emphasis is needed. Denote (recalling that $h^* = h - Eh$, $\underline{h}^* = \underline{h} - E\underline{h}$),

$$Z_n = \frac{1}{\sqrt{n}} \sum_1^n h^*(X_j, D_j)$$

$$Z_n(N,L) = \frac{1}{\sqrt{n}} \sum_1^n \underline{h}^*(X_j,D_j) \ .$$

<u>Proposition 4.4</u>: $E(Z_n - Z_n(N,L))^2 \le c(P(X \in E_N^c))^{1/2}$.

<u>Proof</u>: This follows directly from corollary 2.15.

For the next step, define

$$R_j' = \begin{cases} 0 & \text{if } X_j \in B_\ell, \text{ no other } X_i \in B_\ell \\ \inf\limits_{\substack{i \ne j \\ X_i \in B_\ell}} \|X_i - X_j\| & \text{if } X_j \in B_\ell \end{cases}$$

and redefine $h(x,0) = 0$. Let $D_j' = n^{1/m} R_j'$ and

$$Z_n'(N,L) = \frac{1}{\sqrt{n}} \sum_1^n \underline{h}^*(X_j,D_j') \ .$$

<u>Proposition 4.5</u>: $E(Z_n(N,L) - Z_n'(N,L))^2 \le cne^{-(n-1)\varepsilon_N V(d_N)}$ where $\varepsilon_N > 0$ depends only on N.

<u>Proof</u>:

$$\begin{aligned} E(Z_n(N,L) - Z_n'(N,L))^2 &\le \frac{1}{n} E(\textstyle\sum_j \Delta_j)^2 \\ &\le \textstyle\sum_j E\Delta_j^2 \end{aligned}$$

where

$$\Delta_j = \underline{h}(X_j,D_j) - \underline{h}(X_j,D_j') - E(\underline{h}(X_j,D_j) - \underline{h}(X_j,D_j')) \ .$$

so

$$E(Z_n(N,L) - Z_n'(N,L))^2 \le \sum_j E(\underline{h}(X_j,D_j) - \underline{h}(X_j,D_j'))^2 \ .$$

Now $X_j \in E_N$ and $d(X_j,\tilde{B}) > R_j$ implies $R_j' = R_j$. So

$$\begin{aligned} E(Z_n(N,L) - Z_n'(N,L))^2 &\le 2\|h\|^2 \textstyle\sum_j P(R_j \ne R_j',\ X_j \in E_N) \\ &\le 2\|h\|^2 nP(d(X,\tilde{B}) \le R,\ X \in E_N) \end{aligned}$$

where (X,R) stands for (X_1,R_{1n}) by our usual convention. Now

$$P(R \geq r | X=x) = [1 - F(S(x,r))]^{n-1} .$$

Note that $d(X,\tilde{B}) \leq N\sqrt{m}$ for $X \in E_N$. Now

$$\inf_{x \in C_N} \inf_{0 \leq r \leq \sqrt{m}\,N} [F(S(x,r))/V(r)] = \varepsilon_N > 0$$

since $M(r,x) = F(S(x,r))/V(r)$ is jointly continuous on $[0,\sqrt{m}\,N] \times \bar{C}_N$, where $\bar{C}_N$ is the closure of C_N, and since $M(r,x) > 0$ everywhere in $\bar{C}_N \times [0,\sqrt{m}\,N]$. Therefore

$$P(R \geq d(X,\tilde{B}),\ X \in E_N) \leq \int_{X \in E_N} e^{-(n-1)\varepsilon_N V(d(x,\tilde{B}))} f(x)dx .$$

For $x \in E_N$, $d(x,\tilde{B}) \geq d_N$, so

$$P(R \geq d(X,\tilde{B}),\ X \in E_N) \leq e^{-(n-1)\varepsilon_N V(d_N)}$$

and the proposition follows.

For the next step, put $B_0 = C_N^c$, and denote

$$P(X \in B_\ell) = p_\ell , \quad \ell = 0,1,\ldots,L$$

so $\sum_{\ell=1}^{L} p_\ell = 1$. (Assume that for every ℓ, $p_\ell > 0$, otherwise delete B_ℓ.) Let

$$n_\ell = \#(X_j \in B_\ell)$$

so the $(n_0,\ldots,n_L)$ have a multinomial distribution with parameters $(p_0,\ldots,p_L)$. Consider the following construction: draw numbers $n_0,\ldots,n_L$, $\sum n_\ell = n$ from a multinomial distribution with parameters $(p_0,\ldots,p_L)$. Then put n_ℓ points $X_i^{(\ell)}$, $i = 1,\ldots,n_\ell$ into B_ℓ using the distribution

$$F_\ell(dx) = P(X \in dx | X \in B_\ell) .$$

Denote by P_ℓ the joint distribution of $X_i^{(\ell)}$, $i=1,\ldots,n_\ell$, let $R_i^{(\ell)}$ be the nearest neighbor distance to $X_i^{(\ell)}$ from the other points in B_ℓ, and $D_i^{(\ell)} = n^{1/m}R_i^{(\ell)}$. Put

$$T_\ell = \begin{cases} \sum_{i=1}^{n_\ell} \underset{\sim}{h}(X_i^{(\ell)}, D_i^{(\ell)}), & n_\ell > 1 \\ 0 & , \; n_\ell \le 1 \end{cases}$$

Then

$$\sum_{\ell=1}^{L} T_\ell = \sum_{j=1}^{n} \underset{\sim}{h}(X_j, D_j')$$

<u>Proposition 4.6</u>: There are constants $\gamma_{n,\ell}$, $\ell = 1,\ldots,L$ such that $\gamma_{n,\ell} \to \gamma_\ell$ and

$$E(E(T_\ell|n_\ell) - ET_\ell - (n_\ell - En_\ell)\gamma_{n,\ell})^2 \le C(\ell) < \infty$$

where $C(\ell)$ is independent of n.

<u>Proof</u>: Define

$$\begin{aligned} W_\ell(r|x,n_\ell) &= P_\ell(n^{1/m}R_1^{(\ell)} > r | X_1^{(\ell)} = x) \\ &= [1 - F_\ell(S(x, rn^{-1/m}))]^{n_\ell - 1} \end{aligned}$$

Note that

$$E(T_\ell|n_\ell) = n_\ell \int \underset{\sim}{h}(x,r) W_\ell(dr|x,n_\ell) F_\ell(dx)$$

Define

$$\begin{aligned} \chi_n(r|x) &= W_\ell(r|x, np_\ell) \\ &= [1 - F_\ell(S(x, rn^{-1/m}))]^{np_\ell - 1} \end{aligned}$$

and suppressing the dependence on L, let

$$\mu_n = (n_\ell - np_\ell)/(np_\ell - 1)$$

Then

$$W_\ell(r|x,n_\ell) = \chi_n^{\mu_n + 1}$$

Then

$$W_\ell(dr|x_1 n_\ell) = \frac{n_\ell - 1}{np_\ell - 1} \chi_n^{\mu_n} \chi_n(dr|x)$$

$$= (\mu_n + 1)\chi_n^{\mu_n} d\chi_n$$

where $d\chi_n \equiv \chi_n(dr|x)$. This is zero for $\mu_n = -1$, so we eliminate this set in the expectations to follow. Writing $n_\ell = (np_\ell - 1)\mu_\ell + np_\ell$ leads to the expression

$$(4.7) \quad E(T_\ell|n_\ell) = np_\ell(1+\mu_n)^2 \int \underline{h}\chi_n^{\mu_n} d\chi_n \, dP_\ell - \mu_n(1+\mu_n) \int \underline{h}\chi_n^{\mu_n} d\mu_n \, dP_\ell \, .$$

The expectation of the square of the second term in (4.7) above is bounded by $C_\ell \|\underline{h}\|^2/n$, and is henceforth ignored.

Next, expand

$$\chi_n^{\mu_n} = 1 + \mu_n \log\chi_n + \frac{\mu_n^2}{2}(\log \chi_n)^2 \chi_n^{\theta\mu_n} \quad ,$$

where $0 \leq \theta \leq 1$, and substitute into the first term of (4.7). We assert that all terms containing a power of μ_n higher than one have squares whose expectations are uniformly bounded in n. For example

$$(np_\ell)^2 E(\mu_n^2 \int \underline{h}(\log \chi_n) d\chi_n \, dP_\ell)^2 \leq (np_\ell)^2 \|h\| E\mu_n^4 \leq C\|h_1\|^2(1-p_\ell)^2$$

and

$$(np_\ell)^2 E\left\{\mu_n^2(1+\mu_n)^2 \int \underline{h}(\log \chi_n)^2 \chi_n^{\theta\mu_n} d\chi_n dP_\ell\right\}^2$$

$$\leq \|\underline{h}\|^2 (np_\ell)^2 E\left\{\mu_n^2(1+\mu_n)^2 \int (\log \chi_n)^2 \chi_n^{\theta\mu_n} d\chi_n dP_\ell\right\}^2$$

$$\leq 2\|\underline{h}\|^2 (np_\ell)^2 [E\{\mu_n^4(1+\mu_n)^{-2}; \; -1 < \mu_n \leq 0\} + E\{\mu_n^4(1+\mu_n^4); \; \mu_n > 0\}]$$

$$\leq C_\ell \|\underline{h}\|^2 \, .$$

Therefore

$$(4.8) \qquad E(T_\ell|n_\ell) = np_\ell \int \underline{h}(1+\mu_n(2+\log \chi_n)) d\chi_n \, dP_\ell + O_2(1)$$

so

$$E(T_\ell | n_\ell) - ET_\ell = np_\ell \mu_n \int \underline{h}(2 + \log \chi_n) d\chi_n \, dP_\ell + O_2(1) \tag{4.9}$$

where $O_2(1)$ in (4.8) and (4.9) denote quantities such that $\sup_n E(O_2(1))^2 < \infty$. Letting the $\gamma_{n,\ell}$ of the proposition be defined by

$$\gamma_{n,\ell} = \frac{np_\ell}{np_\ell - 1} \int \underline{h}(2 + \log \chi_n) d\chi_n \, dP_\ell$$

The proof will be completed by showing that the integral on the right above converges. For x fixed, $\chi_n(r|x)$ is a non-increasing function of r such that for $x \in \mathrm{Int}(B_\ell)$

$$\chi_n(r|x) \longrightarrow e^{-f(x)V(r)} = \chi_0(r|x) \ .$$

Since $\underline{h}(x,r)$ is a.s. continuous with respect to $d\chi_0 \, dP_\ell$, then

$$\int \underline{h} \, d\chi_n \, dP_\ell \longrightarrow \int \underline{h} \, d\chi_0 \, dP_\ell \ .$$

Now let

$$\tilde{\chi}_n(r|x) = (1 - \log \chi_n(r|x))\chi_n(r|x)$$

so that

$$\tilde{\chi}_n(dr|x) = -(\log \chi_n(r|x))\chi_n(dr|x) \ .$$

For $x \in \mathrm{Int}(B_\ell)$

$$\tilde{\chi}_n(r|x) \longrightarrow (1 + f(x)V(r))e^{-f(x)V(r)} = \tilde{\chi}_0(r|x)$$

and so

$$\int \underline{h}(\log \chi_n) d\chi_n \, dP_\ell \longrightarrow -\int \underline{h} \, d\chi_0 \, dP_\ell \ . \tag{4.10}$$

Proposition 4.11: $\frac{1}{\sqrt{n}} \sum_{\ell=1}^{L} [E(T_\ell | n_\ell) - E(T_\ell)] \xrightarrow{\mathcal{L}} N(0, \sigma^2_{N,L})$

where

$$\sigma^2_{N,L} = \sum_\ell \gamma^2_\ell p_\ell - (\sum \gamma_\ell p_\ell)^2 .$$

Moreover, $n^{-1} E(\sum_{\ell=1}^{L} [E(T_\ell | n_\ell) - E(T_\ell)]^2 \to \sigma^2_{N,L}$.

Proof: Clear from the preceding proposition.

It is useful to recall the dependence of parameters on N and L at this point.

Proposition 4.12: Let

$$U_n = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{L} (T_\ell - E(T_\ell | n_\ell)) . \tag{4.13}$$

Then there is a constant $s^2_{N,L} < \infty$ such that

$$E(U_n^2 | n_1, \ldots, n_L) \xrightarrow[L_1]{a.s.} s^2_{N,L} .$$

Proof: Given $\underline{n} = n_1, \ldots, n_L$, the terms in the sum for U_n are independent. Thus

$$E(U_n^2 | n_1, \ldots, n_L) = \frac{1}{n} \sum_\ell \mathrm{Var}(T_\ell | n_\ell) ,$$

and

$$\begin{aligned} \mathrm{Var}(T_\ell | n_\ell) = {} & n_\ell \mathrm{Var}(\underline{h}(X_1^{(\ell)}, D_1^{(\ell)}) | n_\ell) \\ & + n_\ell (n_\ell - 1) \mathrm{Cov}(\underline{h}(X_1^{(\ell)}, D_1^{(\ell)}), \underline{h}(X_2^{(\ell)}, D_2^{(\ell)}) | n_\ell) \end{aligned}$$

it is then sufficient to show that

$$\mathrm{Var}(\underline{h}(X_1^{(\ell)}, D_1^{(\ell)}) | n_\ell) \xrightarrow[L_1]{a.s.} \text{constant}$$

$$n \, \mathrm{Cov}(\underline{h}(X_1^{(\ell)}, D_1^{(\ell)}), \underline{h}(X_2^{(\ell)}, D_2^{(\ell)}) | n_\ell) \xrightarrow[L_1]{a.s.} \text{constant} .$$

This result can be gotten through a simple modification of propositions 3.7 and 3.10.

Now we are ready for the final steps. We can write

$$(4.14) \qquad Z_n'(N,L) \overset{\mathcal{D}}{=} U_n + V_n ,$$

with U_n defined in (4.13) and

$$V_n = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{L} [E(T_\ell | n_\ell) - ET_\ell] .$$

By $\overset{\mathcal{D}}{=}$ we mean equality in distribution when U_n and V_n have the joint distribution we have implicitly given them. Denote $e_N^2 = P(X \in E_N^c)$.

<u>Proposition 4.15</u>: If $\sigma^2 = \lim_n \mathrm{Var}(Z_n)$, then

$$|\sigma^2 - (s_{N,L}^2 + \sigma_{N,L}^2)| \le ce_N + 2\sigma\sqrt{ce_N} .$$

<u>Proof</u>: By propositions 4.4 and 4.5

$$(4.16) \qquad \overline{\lim_n}\, E(Z_n - Z_n'(N,L))^2 \le ce_N .$$

Use the inequality

$$(4.17) \qquad |EZ_n^2 - EZ_n'^2(N,L)| \le E|Z_n - Z_n'(N,L)|^2 + 2\sqrt{E(Z_n)^2 E(Z_n - Z_n'(N,L))^2}$$

and take $n \to \infty$ to get the result.

<u>Proposition 4.18</u>: Let $\alpha = \sqrt{\max_\ell p_\ell}$ and take $|t|^3 \le \alpha^{-1}$. Note that α depends on both N and L. Let $g_n(t;N,L)$ denote the characteristic function of $Z_n'(N,L)$. Then

$$\overline{\lim_n}\, |g_n(t;N,L) - e^{-(\sigma_{N,L}^2 + s_{N,L}^2)t^2/2}| \le c\alpha|t|^3 .$$

<u>Proof</u>: $$g_n(t;N,L) = E\, e^{it(U_n+V_n)}$$
$$= E(e^{itV_n}E(e^{itU_n}|\underline{n}))\ , \quad \underline{n} = (n_0,\ldots,n_L)\ .$$

Given $\underline{n}$, $U_n = \sum_1^L A_\ell$, with the A_ℓ independent and having the conditional distribution of $T_\ell - E(T_\ell|n_\ell)$ given n_ℓ. Hence

$$E(e^{itU_n}|\underline{n}) = \Pi f_\ell(t)\ , \qquad f_\ell(t) = E(e^{itA_\ell}|n_\ell)\ .$$

Applying corollary 2.3 to A_ℓ,

$$E(A_\ell^2|n_\ell) \leq c_1(n_\ell/n)\ , \quad E(|A_\ell^3|\,|n_\ell) \leq c_2(n_\ell/n)^{3/2}$$

where c_k will denote constants depending only on m, $\|h\|$, and θ_k will be quantities such that $|\theta_k| \leq 1$. Then

$$|1 - f_\ell(t)| \leq \frac{t^2}{2}E(A_\ell^2|n_\ell) \leq (c_1/2)t^2(n_\ell/n)$$
$$\left|f_\ell(t) - 1 + \frac{t^2}{2}E(A_\ell^2|n_\ell)\right| \leq c_2|t|^3(n_\ell/n)^{3/2}\ .$$

Temporarily restrict t to the range $|t|\alpha \leq c_1^{-1/2}/2$. Define

$$B_n = \{\max_\ell(n_\ell/n) \leq 2\max_\ell p_\ell\}\ .$$

On B_n, $|1 - f_\ell(t)| \leq 1/4$, hence

$$\log f_\ell(t) = \log[1 - (1 - f_\ell(t))]$$
$$= -\frac{t^2}{2}E(A_\ell^2|n_\ell) + \theta_1 c_2|t^3|(n_\ell/n)^{3/2} + \theta_2 c_3 t^4(n_\ell/n)^2\ .$$

So

$$\Pi f_\ell(t) = \exp(-\frac{t^2}{2}\sum_\ell E(A_\ell^2|n_\ell) + \Delta_n)$$

where, since $|t^3|\alpha \leq 1$

$$|\Delta_n| \leq c_2|t^3| \sum_\ell (n_\ell/n)^{3/2} + c_3t^4 \sum (n_\ell/n)^2$$

$$\leq c_2|t^3|\alpha + c_3|t^4|\alpha^2 \leq c_4|t|^3\alpha \ .$$

Therefore

$$|e^{\Delta_n}-1| \leq c_5|t|^3\alpha$$

and so, denoting $\beta_n^2 = E(U_n^2|\underline{n})$

$$|\Pi f_\ell(t) - e^{-\beta_n^2t^2/2}| \leq c_5|t|^3\alpha$$

holds on B_n for all t such that $|t^3| \leq \alpha^{-1}$, and $|t|\alpha \leq c_1^{-1/2}/2$. Write

$$g_n(t;N,L) = E(I(B_n)e^{it(U_n+V_n)}) + E(I(B_n^c)e^{it(U_n+V_n)}) \ .$$

Since $P(B_n^c) \rightarrow 0$, the second term goes to zero, so

$$\overline{\lim} \ |g_n(t;N,L) - Ee^{itV_n-\beta_n^2t^2/2}| \leq c_5|t^3|\alpha \ .$$

Combining this with propositions 4.11 and 4.12

$$\overline{\lim} \ |g_n(t;N,L) - e^{-(s_{N,L}^2+\sigma_{N,L}^2)t^2/2}| \leq c_5|t^3|\alpha \ .$$

To complete the proof we need only remove the restriction $|t|\alpha \leq c_1^{-1/2}/2$. But this can clearly be done by increasing the constant c_5.

The stage is now set for the proof of Theorem 4.1. By (4.16)

$$\overline{\lim}_n \ |g_n(t)-g_n(t;N,L)| \leq \overline{\lim}_n \ E|\exp\{it(Z_n-Z_n'(N,L)\} - 1| \leq |t|\sqrt{ce_N} \ ,$$

where $g_n(t)$ is the characteristic function of Z_n. So, by proposition 4.18,

$$\overline{\lim}_n \ |g_n(t) - \exp\{-(s_{N,L}^2+\sigma_{N,L}^2)\frac{t^2}{2}\}| \leq c(|t|^3\alpha + |t|\sqrt{e_N}) \tag{4.19}$$

for $|t|^3\alpha \leq 1$. Now let $N \to \infty$, $L \to \infty$ in such a way that $\alpha \to 0$ and $e_N \to 0$. By proposition 4.15, if $e_N \to 0$, uniformly in L,

$$\lim_N (s^2_{N,L} + \sigma^2_{N,L}) = \sigma^2 .$$

Since the restriction $|t|^3\alpha \leq 1$ is satisfied eventually for any fixed t, as $\alpha \to 0$ we conclude that, for all t,

$$\lim_n g_n(t) = e^{-\sigma^2 t^2/2}$$

and (4.1) follows since the equality of σ^2 and $\sigma^2(h)$ is derived from the moment convergence theorem 3.5.

By considering linear combinations of h's it is clear how the results can be generalized to provide a multidimensional central limit theorem, and the moment convergence theorem 3.5 can be easily modified to give the limiting form of the covariance matrix.

5. The Process $\hat{H}(t)$ and Goodness-of-Fit

First, a Glivenko-Cantelli type theorem is established for H(t).

Let

$$\text{(5.1)} \qquad \lambda(x) = \begin{cases} \dfrac{f(x)}{g(x)} & ; \; g(x) > 0 \\ \infty & ; \; g(x) = 0 \end{cases}$$

and define a d.f. H by,

$$\text{(5.2)} \qquad H(t) = \begin{cases} Et^{\lambda(X_1)}, & 0 \le t < 1 \\ 1 & , \; t \ge 1 \end{cases} ,$$

and

$$\text{(5.3)} \qquad \alpha = H(1) - H(1-) = P[g(X_1) = 0] \quad .$$

Note that if f=g, then $\alpha=0$ and H is the d.f. of the uniform distribution .

Theorem 5.4: If A(iii) holds, as $n\to\infty$,

$$\text{(5.5)} \qquad \sup_y |\hat{H}(y) - H(y)| \overset{a.s.}{\to} 0 \quad .$$

Proof: We begin by showing,

$$\text{(5.6)} \qquad \hat{H}(y) \to H(y) \text{ a.s. } \forall \; 0 \le y < 1$$

and

$$\text{(5.7)} \qquad \hat{H}(1-) \to 1-\alpha = H(1-), \text{ a.s.} \quad ,$$

To prove (5.6) note that by corollary 2.3,

$$P[|\hat{H}(y) - E\hat{H}(y)| \geq \epsilon] = O(n^{-2})$$

and hence by the Borel-Cantelli lemma,

$$(5.8) \qquad \hat{H}(y) - E\hat{H}(y) \to 0 \quad \text{a.s.} \ \forall\, 0 \leq y < 1 \quad .$$

Assertion (5.6) then follows by using (3.7) to show that $E\hat{H}(y) \to H(y)$. Next (5.7) is an immediate consequence of the S.L.L.N To complete the proof of the theorem, let

$$(5.9) \qquad \hat{H}^*(y) = \begin{cases} \dfrac{\hat{H}(y)}{\hat{H}(1-)} , & 0 \leq y < 1 \\ 1 , & y \geq 1 \end{cases}$$

and define H^* similarly in relation to H. By (5.6) and (5.7) $\hat{H}^*$ converges in law to H^* with probability 1. But H^* is continuous and hence by Polya's theorem,

$$(5.10) \qquad \sup_y |\hat{H}^*(y) - H^*(y)| \xrightarrow{a.s.} 0$$

and (5.5) follows from (5.10) and (5.7).

Define a stochastic process on [0,1] by,

$$(5.11) \qquad Z_n(t) = \sqrt{n}\left(\hat{H}(t) - E\hat{H}(t)\right), \quad 0 \le t \le 1 \quad ,$$

and a corresponding Gaussian process Z with mean 0 whose covariance function $\gamma(s,t)$, $s \le t$, is defined by

$$(5.12) \quad \gamma(s,t) = \int fs^{\lambda}\left(1-\int ft^{\lambda}\right) - \left(\log s\int \lambda s^{\lambda} f \int t^{\lambda} f + \log t\int \lambda t^{\lambda} f \int s^{\lambda} f\right.$$

$$\left. + \log s \log t\int t^{\lambda} f \int s^{\lambda} f\right) + \log s\int \lambda(st)^{\lambda} f + \int \lambda(st)^{\lambda} f \int_{B(s,t)} (7^{\lambda}(s,t,w)-1)\,dw\,dx$$

(We write λ, f for $\lambda(x)$, $f(x)$ etc.)

where

$$B(s,t) = \left\{w: \ r_1 \le \|w\| \le r_1 + r_2\right\} .$$

$$\log 7(s,t,w) = \int_{S(0,r_1) \cap S(w,r_2)} dz$$

where

$$V(r_1) = -\log s$$

$$V(r_2) = -\log t \quad .$$

If $f=g$, then $\gamma(s,t)$, $s \le t$, reduces to

$$(5.13) \qquad \gamma(s,t) = s - st(1 + \log t + \log s \log t) + st\int_{B(s,t)} (\mathcal{T}(s,t,w) - 1)\,dw$$

Clearly the processes $Z_n(\cdot)$ can be identified with probability measures on $D[0,1]$ and it will follow as a consequence of our proof that $Z(\cdot)$ can be as well. In fact, if $\alpha = 0$, $Z(\cdot)$ has a.s. continuous sample functions. Our main result is,

<u>Theorem 5.14</u>: Suppose that A and B hold. Then,

$$Z_n \to Z$$

in the sense of weak convergence in $D[0,1]$ where Z is as above and has a.s. continuous sample functions.

Before giving the proof we state and prove the corollary of greatest interest to us.

Let

$$S_0 = n \int_0^1 (\hat{H}(t) - E\hat{H}(t))^2 \, dt$$

$$S_1 = n \int_0^1 (\hat{H}(t) - E\hat{H}(t))^2 \, d\hat{H}(t) = \sum_{j=1}^{n} \left((E\hat{H})(W_{(j)}) - \frac{j}{n}\right)^2$$

<u>Corollary 5.15</u>: If f=g and A holds, both S_0 and S_1 tend in law to $\int_0^1 Z^2(t)\,dt$ where Z has covariance function (5.13).

The corollary is, for S_0, an immediate consequence of Theorem 5.2. By writing

$$S_1 = \int_0^1 Z_n^2 (\hat{H}^{-1}(t))\, dt$$

we see that the corollary follows in this case from Theorems 5.1 and 5.2.

Notes: 1) The theorem can be extended to the case $\alpha > 0$ by a conditioning argument as in Section 2. Of course the Z process is then continuous only on [0,1) and has a jump at 1.

2) It is not possible in Theorem 5.1 to replace $E\hat{H}$ in the definition of Z_n by H. Although $E\hat{H}(t) \to H(t)$, the difference is of the order of $n^{\frac{-2}{m}}$ and will not be negligible for $m > 3$.

Proof of Theorem 5.14: We begin by establishing the tightness of the Z_n sequence using the 4th moment bound proven in Section 2. Let $R_1, \ldots, R_n$ be as in Section 2 and recall that

$$D_i = n^{\frac{1}{m}} R_i, \; i=1,\ldots,n \quad .$$

Lemma 5.16: If A(iii) and B hold, the sequence of processes $\{Z_n\}$ is tight in D[0,1] and any weak limit point is in C[0,1].

<u>Proof</u>: We use a device due to Shorack (1973).

Note that:

$$Z_n(t) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \left(I\left(g(X_i)D_i^m < \frac{-\log t}{K_m}\right) - P\left(g(X_i)D_i^m < \frac{-\log t}{K_m}\right)\right)$$

where K_m is the volume of the unit sphere in E^m. Let

$$Q_n(t) = G_n\left(\frac{-\log t}{K_m}\right) \quad .$$

where G_n is given in corollary 2.5. Note that by B and the dominated convergence theorem G_n is continuous. For given $\delta > 0$, let $t_1 < \ldots < t_K$ be such that,

$$Q_n(t_i) = \frac{i\delta}{\sqrt{n}}, \quad 1 \le i \le K$$

where $$\frac{K\delta}{\sqrt{n}} \le 1 < (K+1)\frac{\delta}{\sqrt{n}} \quad .$$

Let

$$Z_n^*(t) = Z_n(t_i) + \frac{\sqrt{n}}{\delta}(Q_n(t) - Q_n(t_i)(Z_n(t_{i+1}) - Z_n(t_i))$$

$$\text{for } t_i \le t < t_{i+1} \quad , \quad 0 \le i \le K,\ t_0 = 0,\ t_{K+1} = 1$$

Note that

$$Z_n^*(0) = Z_n^*(1) = 0 \quad .$$

An elementary application of corollary 2.5 shows that,

$$E(Z_n^*(t) - Z_n^*(s))^4 \le M(Q_n(t) - Q_n(s))^2, \ \underline{\text{all}}\ s,\ t \tag{5.17}$$

where M depends on δ but is independent of n. Since, under A(iii) and B, dominated convergence implies that for each y,

$$G_n(y) \to \int f(x)\left(1-\exp\left\{\frac{-1}{2}\,\frac{f(x)K_m\, y}{g(x)}\right\}\right) dx$$

a continuous probability distribution; it follows from a slight modification of Billingsley ((1968), Theorems 12.3 and 12.4) that $\{Z_n^*\}$ is tight and that all limit points of $\{Z_n^*\}$ are in C[0,1].

Next note that

$$(5.18)\quad \sup_t |Z_n(t)-Z_n^*(t)| \leq \max\Big\{\sup\{|Z_n(t)-Z_n(t_i)|:\ t_i \leq t < t_{i+1}\}$$

$$+ \frac{\sqrt{n}}{\delta}\left(\sup\{|Q_n(t)-Q_n(t_i)|:\ t_i \leq t < t_{i+1}\}\right)|Z_n(t_{i+1})-Z_n(t_i)|:\ 0 \leq i \leq K\Big\}$$

$$\leq \max\Big\{ \left[|Z_n(t_{i+1})-Z_n(t_i)| + \sqrt{n}(E\hat{H}_n(t_{i+1})-E\hat{H}_n(t_i))\right]$$

$$+ |Z_n(t_{i+1})-Z_n(t_i)| \ :\ 0 \leq i \leq K\Big\}$$

using the monotonicity of $\hat{H}_n(\cdot)$, $E\hat{H}_n(\cdot)$, $Q_n(\cdot)$.

Next note that integrating (2.8) for j=0, implies that for C independent of n, δ,

$$\sqrt{n}\ E(\hat{H}_n(t_{i+1})-\hat{H}_n(t_i)) \leq C\ \sqrt{n}(Q_n(t_{i+1})-Q_n(t_i)) \leq C\delta \quad .$$

Hence,

$$(5.19)\qquad \sup_t |Z_n(t)-Z_n^*(t)| \leq 2\ \max\{|Z_n^*(t_{i+1})-Z_n^*(t_i)|\ :\ 0 \leq i \leq K\} + C\delta$$

But in view of (5.17), some elementary inequalities give

$$(5.20) \quad P[\max\{|Z_n^*(t_{i+1})-Z_n^*(t_i)|: 0 \le i \le K\} \ge \varepsilon]$$
$$\le \varepsilon^{-4} M \sum_{i=0}^{K} (Q_n(t_{i+1})-Q_n(t_i))^2 \le M \frac{\delta}{\sqrt{n}} \to 0 .$$

By (5.18)-(5.20) for each $\delta > 0$, C independent of δ

$$(5.21) \qquad P[\sup_t |Z_n(t)-Z_n^*(t)| > 2C\delta] \to 0 .$$

Since $\{Z_n^*\}$ is tight for each δ, (5.21) implies tightness of $\{Z_n\}$ and a.s. continuity of all limit points. (See, for example, Theorem 4.2 of Billingsley (1968). Note that the dependence of Z_n^* on δ is immaterial.)

Asymptotic normality of $(Z_n(t_1),\ldots,Z_n(t_n))$ follows from the representation given in the introduction,

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h^*(X_i, D_i)$$

with

$$h(x,d) = I(\exp\{-g(x)V(d)\} \le t)$$

and the multivariate extension of theorem 4.1. Similarly the formulae (5.11) and (5.12) for $\gamma(s,t)$ may be obtained after tedious calculations from the appropriate straightforward generalizations of proposition 3.10.

As an immediate consequence of theorem 5.4 and corollary 5.15 we have

<u>Theorem 5.22</u>: The tests which reject when $S_1 \ge c(\alpha)$ where

$$P_g\{\int_0^1 Z^2(t)dt \ge c(\alpha)\} = \alpha ,$$

asymptotically have level α for H: $f = g$ and are consistant against all $f \ne g$ which satisfy A and B.

Proof: That the tests have level α is immediate from corollary 5.15. We check consistency for S_0.

Note first that if $f \neq g$

$$\int_0^1 (H(t) - t)^2 dt > 0 . \tag{5.23}$$

If not, since $H(e^{-s})$ is the Laplace transform of $\lambda(X_1)$ and equals e^{-s} a.e., then $P_f[\lambda(X_1) = 1] = 1$, implying $f = g$ a.e. Write

$$S_0 = \int_0^1 Z_n^2(t)dt + 2\sqrt{n}\int_0^1 Z_n(t)(E_f\hat{H}(t) - E_g\hat{H}(t))dt + n\int_0^1 (E_f\hat{H}(t) - E_g\hat{H}(t))^2 dt .$$

Then

$$\int_0^1 Z_n^2(t)dt = O_p(1)$$

$$\sqrt{n}\int_0^1 Z_n(t)(E_f\hat{H}(t) - E_g\hat{H}(t))dt = O_p(\sqrt{n})$$

$$n\int_0^1 (E_f\hat{H}(t) - E_g\hat{H}(t))^2 dt \sim n\int_0^1 (H(t)-t)^2 dt = O(n)$$

by (5.23). Therefore,

$$S_0 \xrightarrow{P} \infty$$

and consistency follows.

Note: In his thesis M. Schilling (1979) has made a far reaching investigation of the power of this and related tests against contiguous alternatives, has constructed tables of the asymptotic null distribution of S_0 for $m = 1$ and ∞ and has studied the efficiency of the large m and n approximation through simulation.

APPENDIX

In this appendix we give the statements and proofs of several lemmas of a technical or computational nature which are used in the previous sections. We begin with a key lemma due to Stone (1977).

Lemma S: For each m and norm $||\cdot||$ there exists $\alpha(m) < \infty$ such that it is possible to write R^m as the union of $\alpha(m)$ disjoint cones $C_1,\ldots,C_\alpha$ with 0 as their common peak such that if

$$x,\ y \in C_j, x,y \neq 0,\ \text{then}\ \ ||x-y|| < \max(\,||x||\,,\,||y||\,),\ j=1,\ldots,\alpha(m) \quad .$$

The following straightforward modification of Stone's argument shows that the lemma is valid for any norm.

Proof: By compactness of the surface of the unit sphere $\partial S(0,1)$ we can find $\tilde{C}_1,\ldots,\tilde{C}_{\alpha(m)}$ disjoint sets such that,

$$\text{(i)} \quad \cup_{j=1}^{\alpha(m)} \tilde{C}_j = \partial S(0,1)$$

$$\text{(ii)} \quad x,\ y \in \tilde{C}_j \Longrightarrow ||x-y|| < 1 \quad .$$

Let

$$C_j = \{\lambda x : x \varepsilon\ \tilde{C}_j,\ \lambda \geq 0\},\ j=1,\ldots,\alpha(m) \quad .$$

Suppose $x = \lambda\tilde{x}$, $y = \gamma\tilde{y}$, $\tilde{x}$, $\tilde{y} \varepsilon \tilde{C}_j$. Suppose w.l.o.g. $\lambda \leq \gamma$. Then,

$$||x-y|| = \gamma\,||\frac{\lambda}{\gamma}\tilde{x}-\tilde{y}|| \leq \gamma\left\{\left(1-\frac{\lambda}{\gamma}\right)||\tilde{y}|| + \frac{\lambda}{\gamma}\,||\tilde{x}-\tilde{y}||\right\} < ||y||$$

The following are easy corollaries of lemma S.

<u>Corollary S1</u>: For any set of n distinct points, $x_1,\ldots,x_n$ in R^m, x_1 can be the nearest neighbor of at most $\alpha(m)$ points.

<u>Corollary S2</u>: If $C_1,\ldots,C_{\alpha(m)}$ are as in lemma S, y_0 is arbitrary, $x \in C_j+y_0$, then

$$S(x,\|x-y_0\|) \supset S(y_0,\|x-y_0\|) \cap (C_j+y_0)\ .$$

The following consequence of S2 is needed for the proof of lemma A2 but is of independent interest.

<u>Theorem A1</u>: Let Y be a random m vector with distribution G, density g, and let y_0 be a fixed point,

$$Q = G(S(Y,||Y-y_0||))$$

Then,

(A.2) $$P[Q \leq q] \leq \alpha(m)\, q,\ 0 \leq q \leq 1\quad .$$

<u>Proof</u>: First let $y_0 = 0$ and let G_j be the conditional distribution of $Y|Y \varepsilon C_j$ and $p_j = G(C_j)$, where the C_j are given by corollary S2. Then,

(A.3) $$P[Q \leq q] = \sum_j\left\{p_j\ P[Q \leq q | Y \varepsilon C_j] : p_j > 0\right\}$$

But $Y \in C_j$ implies by corollary S2 that

$$G(S(Y, \|Y\|)) \geq p_j G_j(S(0, \|Y\|) \cap C_j) .$$

Hence, for $p_j > 0$.

$$P[Q \leq q | Y \in C_j] \leq P[G_j(S(0, \|Y\|)) \leq q/p_j | Y \in C_j] = \frac{q}{p_j} \tag{A.4}$$

since, given $Y \in C_j$, $G_j(S(0, \|Y\|))$ has a uniform distribution on $(0,1)$. (A.2) and (A.3) imply (A.1) if $y_0 = 0$. For the general case shift everything by y_0 and apply corollary S2 in full generality.

<u>Corollary A5</u>: If Q is as in theorem A.1, $r \geq 0$

$$E(1-Q)^r Q \leq M(r+1)^{-2}$$

where M depends only on m.

<u>Proof</u>: Since $0 \leq Q \leq 1$ we may w.l.o.g. take $r \geq 2$. By integration by parts

$$\begin{aligned} E(1-Q)^r Q &= \int_0^1 P[Q \leq q]\{-(1-q)^r + rq(1-q)^{r-1} dq\} \\ &\leq \alpha(m) r \int_0^1 q^2 (1-q)^{r-1} dq \\ &\leq r(r-1)^{-3} \alpha(m) \int_0^{r-1} w^2 \left(1 - \frac{w}{r-1}\right)^{r-1} dw \\ &\leq 2\alpha(m) r(r-1)^{-3} \\ &\leq M(r+1)^{-2} \end{aligned}$$

We proceed to lemmas A6 and A10.

Lemma A6: Let

$$F_{i1} = [X_i \neq \tilde{X}_i]$$

$$F_{i2} = [X_i = \tilde{X}_i,\ R_i \neq \tilde{R}_i]$$

$$F_{i3} = [J_i = 2 \text{ or } \tilde{J}_i = 2] \quad .$$

Then

$$\text{(A.7)} \qquad P[F_{1j}] \leq M\left(\frac{r}{n} + F(S)\right),\ \forall j$$

$$\text{(A.8)} \qquad P[F_{1j} \cap F_{1k}] \leq M\left(\frac{r^2}{n^2} + F^2(S)\right),\ \forall j \neq k \quad .$$

Proof: All these estimates follow by symmetry arguments as in the proof of lemma 2.27. We prove one of the estimates of (A.8) as an example. Note that we may without loss of generality take $r \leq n/4$ (say). Then

$$\text{(A.9)} \quad P[F_{12} \cap F_{13}] \leq [(n-r)(n-r-1)]^{-1} E[\sum_{i=1}^{n-r} I(F_{i2}) \sum_{\substack{k=1 \\ k \neq i}}^{n-r} (I(J_i = k) + I(\tilde{J}_i = k))]$$
$$\leq 8\alpha(m) n^{-2} E(N+r)$$

by corollary S1. But

$$8\alpha(m) n^{-2} E(N+r) \leq \frac{M}{n}\left(\frac{r}{n} + F(S)\right) \leq M\left(\frac{r^2}{n^2} + F^2(S)\right) \quad .$$

Clearly the bounds (A.7) and (A.8) are overestimates in this case. We have written the lemma in this way for compactness.

Lemma A10: With the same definitions for $j = 1,2$,

(A.11) $$E\, I(F_{1j})\, \frac{N}{n} \leq M\left(\frac{r^2}{n^2} + F^2(S)\right)$$

(A.12) $$E\, I(F_{1j})\, F(S_1) \leq M\left(\frac{r^2}{n^2} + F^2(S)\right)$$

(A.13) $$E\, I(F_{1j})\, F(\tilde{S}_1) \leq M\left(\frac{r^2}{n^2} + F^2(S)\right)$$

Proof: a) $j = 1$

$$E\, I(F_{11})\, \frac{N}{n} = F(S)\left(1 + \left(1 - \frac{r-1}{n}\right) F(S)\right)$$

$$E\, I(F_{11})\, F(S_1) = P(F_{11})\, EF(S_1) = \frac{F(S)}{n} \quad .$$

Let

$$R_i^* = \min\{||\tilde{X}_i - \tilde{X}_j|| : 1 \leq j \leq n-r,\ j \neq i\} \quad .$$

Then,

$$E\, I(F_{11})F(\tilde{S}_1) \leq E\, I(F_{11})F(S(\tilde{X}_1, R_1^*))$$

$$\leq (n-r)^{-1}\, F(S)(1-F(S)) + F^2(S) \quad .$$

The bounds (A.11 - A.13) are immediate for $r \leq n/4$ and trivial (for large enough M) for $r > n/4$.

b) $j = 2$

$$E\, I(F_{12}) \frac{N}{n} = \left(1 - \frac{r-1}{n}\right) P[F_{12} \cap F_{21}] \le 2\alpha(m) \frac{EN(N+r)}{n(n-r-2)}$$

$$\le M\left(\frac{r}{n} F(S) + F^2(S)\right)$$

for $r \le n/4$ and (A.11) follows.

To prove (A.12) begin by writing,

$$\text{(A.14)} \qquad E\, I(F_{12}) F(S_1) \le E\, I(X_1 = \tilde{X}_1, R_1 < \tilde{R}_1) F(S_1)$$

$$+ E\, I(X_1 = \tilde{X}_1, R_{10} > R_{1c}) F(S_1) + \sum_{j=1}^{r} E\, I(X_1 = \tilde{X}_1, R_{10} > \|X_1 - x_j\|) F(S_1)$$

where,

$$R_{10} = \min\{\|\tilde{X}_j - \tilde{X}_1\| : X_j = \tilde{X}_j,\ j \ne 1,\ 1 \le j \le n-r\}$$

$$R_{1c} = \min\{\|\tilde{X}_j - \tilde{X}_1\| : X_j \ne \tilde{X}_j,\ j \ne 1,\ 1 \le j \le n-r\}$$

Then, we bound

$$\text{(A.15)} \qquad E\, I(X_1 = \tilde{X}_1, R_1 < \tilde{R}_1) F(S_1) \le E\, I(X_{J_1} \ne \tilde{X}_{J_1}) F(S_1) = n^{-1} F(S)$$

Next,

$$\text{(A.16)} \qquad E\, I(X_1 = \tilde{X}_1, R_{10} > R_{1c}) F(S_1)$$

$$\leq E\{P[F_S(S(\tilde{X}_1, R_{10})) > F_S(S(\tilde{X}_1, R_{1c})) | N, \tilde{X}_1, R_{1c}, X_1 = \tilde{X}_1] F_S(S(\tilde{X}_1, R_{1c})) I(X_1 = \tilde{X}_1)\}$$

$$= E[(1 - F_S(S(\tilde{X}_1, R_{1c})))^{K-1} F_S(S(\tilde{X}_1, R_{1c})) I(X_1 = \tilde{X}_1)]$$

where $K = n-r-N$

$$\leq E\, N \int_0^1 (1-w)^{n-r-2} w\, dw = [(n-r)(n-r-1)]^{-1} EN$$

$$\leq M \frac{r}{n} F(S)$$

for $r \leq n/4$.

The next to last inequality follows since, given $X_1 = \tilde{X}_1$ and N, $F_S(\tilde{X}_1, R_{1c})$ is distributed as the minimum of N uniform (0,1) variables. Finally, arguing as above,

$$\text{(A.17)} \qquad E\, I(X_1 = \tilde{X}_1, R_{10} > ||X_1 - x_j||) F(S_1)$$

$$\leq E(1 - F_S(S(\tilde{X}_1, ||\tilde{X}_1 - x_j||))^{K-1} F_S(S(\tilde{X}_1, ||\tilde{X}_1 - x_j||) I(X_1 = \tilde{X}_1)$$

Given $X_1 = \tilde{X}_1$, we can apply corollary A.1 noting that $F_S(S(\tilde{X}_1, ||\tilde{X}_1 - x_j||))$ has the distribution of Q with $G = F_S$, $x_j = y_0$. Since conditionally K-1 has a binomial (n-r-1, 1-F(S)) distribution, we obtain as a bound for (A.17),

$$ME(K^{-2}|X_1=\tilde{X}_1) \leq \frac{3}{2} M(1-F(S))^{-2}(n-r)^{-2} \quad . \tag{A.18}$$

Therefore, we obtain

$$\sum_{j=1}^{r} EI(X_1=\tilde{X}_1, R_{10} > ||X_1 - x_j||)F(S_1) \leq M\left(\frac{r^2}{n^2} + F(S)\right) \tag{A.19}$$

for $\qquad r \leq \frac{n}{4}$, $F(S) \leq \frac{1}{4}$ (say).

Combining (A.15), (A.16) and (A.17) we obtain (A.12) for $j=2$, since the restrictions on r and F can be absorbed into M for the final bound. Finally,

$$E\ I(F_{12})\tilde{S}_1 \leq E\ I(X_1=\tilde{X}_1, \tilde{R}_1 < R_1)F(S_1) + E\ I(X_1=\tilde{X}_1, X_{J_1} \neq \tilde{X}_{J_1})F(S(\tilde{X}_1, R_{10})) \quad . \tag{A.20}$$

The first term in (A.20) has been bounded in (A.14) and (A.19). The second is bounded as in (A.15) by

$$F(S)\ E(\tfrac{1}{K}|X_1=\tilde{X}_1) \leq M\ F(S)\ \frac{r}{n}\ ,\ F(S) \leq \frac{1}{4} \quad ,$$

$r \leq n/4$. (A.13) follows for $j=2$ and the lemma is proved.

REFERENCES:

Billingsley, P. (1968) Convergence of Probability Measures, Wiley.

Hoeffding, W. (1963) Probability Inequalities for Sums of Bounded Random Variables, J. Amer. Statist. Assoc., 58, 13-30.

Loftsgaarden, D. O. and Quesenberry, C. P. (1965) A Nonparametric Estimate of a Multivariate Density Function, Ann. Math. Statist., 36, 1049-1051.

Morris, C. (1975), Central Limit Theorems for Multinomial Sums, Ann. Statist., 3, 165-188.

Rogers, W. (1977) Thesis, Stanford University.

Schilling, M. (1979) Thesis, University of California, Berkeley.

Shorack, G. (1973) Convergence of Reduced Empirical and Quantile Processes, etc., Ann. Statist., 1, 146-152.

Stone, C. (1977) Consistent Nonparametric Regression, Ann. Statist., 5, 595-645.